Elucidation of Nonadditive Effects in Protein–Ligand Binding Energies: Thrombin as a Case Study

Elucidation of Non-Additive effects in Protein-Ligand Binding Energies: Thrombin as a Case Study.

*Gaetano Calabrò†‖, Christopher J. Woods§, Francis Powlesland†‖, Antonia S. J. S. Mey,† Adrian J. Mulholland‡, Julien Michel †**

† EaStCHEM School of Chemistry, University of Edinburgh, Joseph Black Building, King's Buildings, Edinburgh EH9 3JJ, UK

§ BrisSynBio, Life Sciences Building, Tyndall Avenue, Bristol, BS8 1TQ, UK

‡ School of Chemistry, Cantock's close, University of Bristol, BS8 1TS, UK

Abstract: Accurate predictions of free energies of binding of ligands to proteins are challenging partly because of the non-additivity of protein-ligand interactions, i.e. the free energy of binding is the sum of numerous enthalpic and entropic contributions that cannot be separated into functional group contributions. In principle molecular simulations methodologies that compute free energies of binding do capture non-additivity of protein-ligand interactions, but efficient protocols are necessary to compute well-converged free energies of binding that clearly resolve non-additive effects. To this end an efficient GPU-accelerated implementation of alchemical free energy calculations has been developed and applied to two congeneric series of ligands of the

enzyme thrombin. The results show that accurate binding affinities are computed across the two congeneric series and positive coupling between non polar R^1 substituents and a $X=NH_3^+$ substituent is reproduced, albeit with a weaker trend than experimentally observed. By contrast a docking methodology completely fails to capture non-additive effects. Further analysis shows that the non-additive effects are partly due to variations in the strength of a hydrogen-bond between the $X=NH_3^+$ ligands family and thrombin residue Gly216. However other partially compensating interactions occur across the entire binding site and no single interaction dictates the magnitude of the non-additive effects for all the analysed protein-ligand complexes.

Introduction

A significant component of pre-clinical drug discovery is the optimization of the non-covalent binding between a disease-involved protein and small organic molecules.[1–3] To this end, the free energy of binding is frequently used to quantify the strength of protein-ligand interactions. Separating the free energy of binding in enthalpic and entropic components indicates that potent ligands may be obtained by minimizing the enthalpy of binding while maximizing the entropy of binding. Binding enthalpies result from a competition between two concomitant effects: Van der Waals (VdW) and electrostatic interactions, and the desolvation of polar groups.[4] VdW interactions are optimised by the shape complementarity between the protein and the ligand, whereas electrostatic interactions are optimised via charge and hydrogen bonding interactions between donor and acceptor atoms in the protein-ligand complex[5]. The desolvation of polar groups reflects the strength of the interactions between the solvent and the protein/ligand, before formation of the protein-ligand complex. A favourable enthalpy change upon binding is an indication of sufficiently strong VdW/electrostatic interactions between the target and the ligand

that compensate for the unfavourable enthalpy change associated with desolvation[5]. Binding entropies arise from two major terms: conformational entropy changes and solvation entropy changes. The former is related to changes in probability distributions of translational, rotational and internal degrees of freedom of the ligand and protein upon complex formation, while the latter depends on differences in the probability distributions of translational and rotational degrees of freedom of the water molecules that solvate the protein, the ligand, and the complex.[5–7]

In this intricate framework, medicinal chemists are usually tasked to identify structural modifications of a ligand that will optimise all the above contributions to the free energy of binding. This is frequently pursued through iterative Structure-Activity Relationship (SAR) studies where individual moieties of a ligand are optimised sequentially.[8,9] A major simplification implicit in this approach is the additivity of the binding free energy, i.e. the assumption that binding free energies can be decomposed into a sum of independent components ascribed to specific parts of the system. Many popular molecular modelling methods also make this assumption, for instance Matched Molecular Pair Analysis (MMPS),[10] Free-Wilson Analysis,[11] scoring function approaches and linear QSAR models.[12]

In general, if the cruel question is: "can free energy be decomposed into a sum of independent components ascribed to specific parts in a system?", then the answer is negative. Free energy is a property of the whole of phase space and its decomposition in components holds only if phase space is divisible into independent parts. This result has been reported several times in the literature.[13–15] Assuming for simplicity a canonical ensemble, the Helmholtz free energy F of a thermodynamic system at temperature T and Hamiltonian H is given by:

$$F = -k_b T \ln\left\langle e^{-\beta H}\right\rangle_{NVT} , \quad (1)$$

where $\beta = 1/k_bT$, k_b is the Boltzmann constant, and the symbol <> denotes a canonical ensemble average. If the Hamiltonian H can be separated into for instance two independent components one can write equation 2:

$$H(q_1, q_2; p_1, p_2) = \tilde{H}_1(q_1; p_1) + \tilde{H}_2(q_2; p_2)\,, \tag{2}$$

where it is implied that q_1, p_1, q_2 and p_2 are uncoupled position and momenta coordinates, then (and only then) it follows that:[13]

$$F = -k_bT \ln\langle e^{-\beta\tilde{H}_1} e^{-\beta\tilde{H}_2}\rangle_{NVT} = -k_bT \ln\langle e^{-\beta\tilde{H}_1}\rangle_{NVT} - k_bT \ln\langle e^{-\beta\tilde{H}_2}\rangle_{NVT} = F_1 + F_2 \tag{3}$$

The extent to which the additivity assumption in eq 3 may be valid in practice is a question that is of central importance in many chemical contexts. Patel et al.[16] investigated 19 different protein targets to examine the extent of non-additive substituent effects on ligands binding affinities and found that only half of the targets exhibited approximately additive behaviour. Thus, while *in theory* it is incorrect to expect that free energies of binding may be decomposed into independent components; *in practice* additivity of protein-ligand interactions may be a good approximation in some instances. It follows that it is desirable to formulate molecular models that may reproduce, and elucidate, non-additive effects in protein-ligand interactions.

A detailed study of non-additivity of functional group contributions to protein-ligand interactions was reported by Baum et al.[17] The work featured isothermal titration calorimetry (ITC) measurements and crystallographic analyses of the enzyme thrombin in complex with different series of congeneric inhibitors. Structurally the ligand binding site in thrombin presents three sub-pockets illustrated in Figure 1A. The ligands may be classified in two series that shall be referred to as series 3 (X=H) and series 5 (X=NH_3^+) for consistency with the notation used by

Baum et al. The amino group of series 5 ligands forms a charged assisted hydrogen-bond with the backbone carbonyl of Gly216. Each series also features variations in the nature of the R^1 group that fills the S1 pocket.[17] Other series were reported, but are not the focus of the present study. Baum et al. demonstrated coupling between the X and R^1 substituents since the standard free energy of binding of series 5 ligands is more favourable than series 3 ligands with equivalent R^1 groups (Figure 1B). Correlation of structural data from X-ray diffracted crystals and calorimetric data suggested that the non-additivity was correlated to changes in hydrogen-bonding distances between the ligands and the backbone of Gly216, changes in residual mobility of the ligands, and changes in packing interactions with S3 pocket residues.

Which of these factors dominate non-additivity in this context remains uncertain owing to the difficulty of disentangling cause and effects from the experimental data. Also given potential pitfalls in relating structural and dynamical measurements from X-ray diffracted crystals to protein-ligand interactions in aqueous conditions,[18] further insights into plausible non-additivity mechanisms is desirable, and this was here sought by means of molecular simulations.

Materials and Methods

Protein structure setup

The crystallographic structure of human thrombin in a complex with a thrombin ligand structurally related to the ligands simulated in this study was downloaded from the PDB databank[19] (PDB code 2ZC9[17]). The protein structure was inspected and prepared for molecular simulations using Maestro.[20] The hirugen chain was removed from the structure. The side chain of Arg75 in the heavy chain (chain H) of thrombin was partially resolved.[19] Missing atoms were added to complete the side-chain in a solvent exposed conformation.[19] The incomplete light

chain was capped before Glu1C with an ACE residue and after Ile14L with an NME residue.[19] The incomplete heavy chain was capped after Gly246 with an NME residue.[19] Missing residues Trp148, Thr149, Ala149A, Asn149B, Val149C, Gly149D, Lys149E in chain H were modeled in the structure using the FALC-Loop web server[21]. Standard protonation states were assumed for protein side-chains[19]. On the basis of visual inspection of hydrogen bonding patterns, His57 and His71 were modeled in their uncharged δ-tautomer. His91, His119 and His230 were modeled in the δ-tautomer[19]. Disulfide bridges were modeled between Cys42-Cys58, Cys1-Cys122, Cys168-Cys182 and Cys191-Cys220[19].

Ligands structure setup

In this investigation all the thrombin inhibitors in the 3 and 5 series reported by Baum et al. study were considered (Figure 1).[17] The ligands in the 3 and 5 series were modeled using the software Maestro,[20] and manually placed inside the binding site in a starting conformation selected according to available crystallographic data reported by Baum et al., and visual inspection. In addition, the amino group X=NH_3^+ for each ligand in the 5 series was protonated for consistency with the *in vitro* assay conditions.

Molecular simulations setup

In order to assemble input files used as starting conformations for subsequent calculations the FESetup software package was used.[22] The following protocol was set in FESetup for the automated preparation of the ligands, protein and complexes input files:

Ligands: The atomic charges were assigned by using Antechamber[23] (as implemented in Amber 11), selecting the AM1-BCC method.[24] The GAFF force field was used for the generation of the remaining force field parameters.[23,25] The Leap software was instructed to solvate ligands in a buffer of TIP3P water molecules,[26] and counter-ions were added for the resulting systems to

yield a net charge of zero.[27] The solvated systems were energy minimized for 100 cycles by using the steepest descent method and equilibrated at 300 K and 1 atm pressure for 10^5 MD steps with a 2 fs time step using the Amber module Sander. During the equilibration stage, the ligands were restrained to their starting positions using a harmonic potential with a force constant of 10 kcal mol^{-1} $Å^{-2}$ and constraining all the hydrogen bonds to their equilibrium distances.

Protein: The Amber ff99SB force field parameters were used to parameterize the protein.[28] The protein was energy minimized in vacuum for 500 cycles of steepest descent method by using the Amber module Sander.

Protein-ligand complexes: The ligands were combined with the thrombin protein model and solvated in a buffer of TIP3P water molecules,[26] and counter ions were also added to neutralize the solution.[27] The complexes were energy minimized for 500 cycles by using the steepest descent method and equilibrated at 300 K and 1 atm pressure for 10^5 MD steps with 2 fs time step via the Amber module Sander. During the equilibration stage, the protein and the ligands were restrained to their starting positions using a harmonic potential with a force constant of 10 kcal•mol^{-1}•$Å^{-2}$ and constraining all the hydrogen bonds to their equilibrium distances.

Protein-morph complexes and solvated morphs: A single topology alchemical free energy calculation protocol was adopted in this study (see below). Thus coordinates and parameters for morph molecules that map between pairs of ligands were generated by the software FESetup.[22] Instructions to download all input files generated by the procedure are given in the Supporting Information.

Molecular simulations protocols

Figure 2 illustrates the strategy used to study non-additive effects in series 3 and 5 respectively. The symbols 5R' and 5R'' denote ligands with X=NH_3^+ and R^1=R' or R'' substituents. Similarly, the symbols 3R' and 3R'' denote ligands with X=H and R^1=R' or R''

substituents. The non-additivity level (*NA*) quantity in the isothermal-isobaric ensemble was defined as follows:

$$NA = \Delta\Delta G_{bind}(5\mathrm{R}' \rightarrow 5\mathrm{R}'') - \Delta\Delta G_{bind}(3\mathrm{R}' \rightarrow 3\mathrm{R}'') \,, \quad (4)$$

where $\Delta\Delta G_{bind}(5\mathrm{R}' \rightarrow 5\mathrm{R}'') -$ and $\Delta\Delta G_{bind}(3\mathrm{R}' \rightarrow 3\mathrm{R}'')$ are the binding free energies of 5R'' and 3R'' relative to 5R' and 3R' respectively. If *NA* equates zero then the nature of the group X does not influence the interactions of R' or R'', otherwise *NA* quantifies the extent of positive or negative coupling between the interactions of X and R'/R''.

Equation (4) shows that in order to compute non-additivity levels it is necessary to calculate the relative binding free energy between the different ligands in the two series. Each binding free energy of a ligand **B** relative to a ligand **A** is the difference of two computed free energy changes given by:

$$\Delta\Delta G_{bind}(\mathbf{A} \rightarrow \mathbf{B}) = \Delta G_{prot}(\mathbf{A} \rightarrow \mathbf{B}) - \Delta G_{wat}(\mathbf{A} \rightarrow \mathbf{B}) \,, \quad (5)$$

where $\Delta G_{prot}(\mathbf{A} \rightarrow \mathbf{B})$ is the computed free energy change for morphing a ligand **A** into a ligand **B** in the solvated protein binding site, and $\Delta G_{wat}(\mathbf{A} \rightarrow \mathbf{B})$ is the computed free energy change for morphing ligand **A** into **B** in an aqueous environment.[29,30] The starting and final ligands used to compute the relative binding affinities in this study are reported in Figure 3. The ligands **3B** and **5B** were selected as reference end-state most of the time because these ligands were the most central in the network of the defined alchemical transformations.

Similarly, it is also possible to define the contribution of components of the potential energy function to non-additivity levels. For instance, for protein-ligand interaction energies:

$$NA_{cont}(U_{PL}) = [\langle U_{PL}(5\mathrm{R}'')\rangle - \langle U_{PL}(5\mathrm{R}')\rangle] - [\langle U_{PL}(3\mathrm{R}'')\rangle - \langle U_{PL}(3\mathrm{R}')\rangle] \,, \quad (6)$$

where U_{PL} is the protein-ligand interaction energy, and the brackets denote ensemble averages.

Equation (4) indicates that non-additivity levels result from differences in two relative free energies of binding. This implies that precise measurements of non-additivity levels require two well converged relative binding free energies. Given the potentially slow convergence of these quantities it was deemed important to achieve high-performance in conformational sampling to produce well-converged free energies. To this end an *ad-hoc* implementation based on the alchemical single topology method was developed to calculate the relative binding free energies of the thrombin ligands. The implementation extended the molecular simulation library Sire[31],

through linking to the OpenMM API.[32] This API is able to implement Molecular Dynamics (MD) algorithms by using CUDA and OpenCL architectures present on the latest Graphic Processing Units (GPUs) and gain significant computational power compared to more conventional approaches.[33–35]

The ligand mutations in both environments were performed by using the single topology method as implemented in the software package Sire release 2270. The coupling parameter λ used to mutate the ligands in both environments was modulated in the λ range [0, 1] where, $\lambda = 0$ maps the morph molecule to ligand **A**, and $\lambda = 1$ maps the morph molecule to ligand **B**. The transformations were performed selecting 16 λ_i values (0.00000, 0.00616, 0.02447, 0.07368, 0.11980, 0.19045, 0.28534, 0.40631, 0.57822, 0.70755, 0.80955, 0.88020, 0.92632, 0.97553, 0.99384 and 1.00000). The λ_i values were chosen to concentrate at the beginning and end of the integration interval to improve numerical stability of the polynomial regression technique used to estimate free energy changes (see below).[36]

In the single-topology implementation done here, the intra- and inter- molecular force field parameters involved in the mutations were linearly interpolated between the starting and final

ligand mutant parameters. In order to circumvent steric clashes at the end points of the simulations a soft core potential was used for non-bonded interactions of atoms that can be created or annihilated.[37] The *n* and *d* parameters that control the rate of softening of Coulombic and Lennard-Jones interactions in the equation 3 of ref 37 were respectively set to 0 and 2.0.

Free energy changes were calculated by using the thermodynamic integration method:[38,39]

$$\Delta G(\mathbf{A} \rightarrow \mathbf{B}) = \int_0^1 \frac{dG}{d\lambda} d\lambda \,, \tag{7}$$

where the integral in eq 7 was numerically estimated by using a polynomial interpolation of seventh-order.[36] The free energy gradients at each selected λ_i value were approximated by using a finite difference approach:[40]

$$\frac{dG}{d\lambda} \simeq \frac{\Delta G(\lambda \rightarrow \lambda + \Delta\lambda) - \Delta G(\lambda \rightarrow \lambda - \Delta\lambda)}{2\Delta\lambda}, \tag{8}$$

where the delta increment was set to $\Delta\lambda = 10^{-3}$. The free energy changes $\Delta G(\lambda \rightarrow \lambda \pm \Delta\lambda)$ were calculated by using the Zwanzig equation given by eq 9:[41]

$$\Delta G(\lambda \rightarrow \lambda \pm \Delta\lambda) = -1/\beta \ln < e^{-\beta(U(\lambda \pm \Delta\lambda) - U(\lambda))} >_\lambda \tag{9}$$

where U is the total system potential energy and $<>_\lambda$ denotes the ensemble average at the selected λ value. Contribution of pressure-volume changes were neglected in eq 9. For the evaluation of free energy gradients at λ=0.0 and λ=1.0 assumptions were made that $\Delta G(\lambda \rightarrow \lambda - \Delta\lambda = -\Delta G\lambda \rightarrow \lambda + \Delta\lambda$ and $\Delta G\lambda \rightarrow \lambda + \Delta\lambda = -\Delta G\lambda \rightarrow \lambda - \Delta\lambda$.

For each λ value the ensemble average was computed by sampling the system using MD as implemented in Sire via the OpenMM API release 5.2. The OpenCL platform was used throughout. In production runs, each window λ was sampled for 10 ns using the NPT ensemble by setting the pressure and the temperature respectively to 1 atm and 300 K. The pressure was

regulated using a Monte Carlo Barostat with an update frequency of 25 MD steps,[42,43] and the Andersen Thermostat was used to keep the temperature constant,[44] selecting a collision frequency of 10 ps^{-1}. The simulations were time evolved by using a Leapfrog-Verlet integrator with a 2 fs time step. All bonds involving hydrogens were constrained to their equilibrium distances. Non-bonded interactions were evaluated by using an atom-based cut off scheme setting the cutoff distance to 10 Å. Long-range electrostatic interactions were calculated by using an atom-based Barker-Watts reaction field,[45] with the medium dielectric constant set to the water dielectric constant ($\varepsilon_{solvent}$ = 78.3).[46] A total of 5×10^4 gradient values were collected over 10 ns simulations for each λ value and each protein-ligand complex calculation was repeated at least three times per ligand. At the beginning of each repeated run the particle velocities were randomly generated accordingly to the Maxwell-Boltzmann distribution at a temperature of 300 K. In order to circumvent steric clashes at the beginning of the production runs due to modifications of the ligand parameters via changes in λ, each starting structure was energy minimized for 500 steps and re-equilibrated. This was done incrementing λ in steps of 0.1 until the desired target value was reached. Each intermediate λ value was equilibrated for 2 ps using a 0.5 fs time step, keeping pressure and temperature set to 1 atm and 300 K. At the end of the re-equilibration stage the time step was re-set to 2 fs. All the production runs were performed on an in-house GPU cluster of nVidia Tesla M2090 graphic card units, and the Blue Crystal GPU cluster of nVidia Tesla K20 graphic card units. Instructions to download sample simulation scripts are provided in the Supporting Information.

B-factors analyses

B-factors for ligand atoms were computed from the simulations in order to quantify ligand mobility in the thrombin binding site. This was done using equation 10 for selected ligand heavy atoms:

$$B = \frac{8\pi^2}{3}\langle \Delta r^2 \rangle \,, \tag{10}$$

where Δr^2 is the mean square of the atomic displacement averaged over a molecular dynamics trajectory, after rigid body alignment against a reference coordinate. The reference coordinate was taken to be the average atomic position recorded along the trajectory. This analysis was performed for the end-states (λ=0 and λ=1) of each free energy calculation and ligand B-factors were computed by averaging results from all individual trajectories.

Correlation errors analysis

The extent to which uncertainties in measured and computed quantities influence statistical metrics used to evaluate correlations between experiment and measurements should be carefully assessed to determine the relevance of the results. This was done here using an error analysis procedure described below.[47,48]

The experimental e_i and the predicted p_i relative binding free energies related to the ligand $i \in \{1, \dots, N\}$ (N is the total number of the ligands) can be represented as the sets $\{\dots, e_i \pm \Delta e_i, \dots\}$ and $\{\dots, p_i \pm \Delta p_i, \dots\}$, where Δe_i and Δp_i are respectively the experimental and predicted statistical uncertainties. It is assumed that the measured relative binding affinities are normally distributed and, therefore, for each data point e_i and p_i it is possible to generate new data points $\widetilde{e_i}$ and $\widetilde{p_i}$ drawing from a normal distribution $\widetilde{e_i} \sim N(\mu = e_i, \sigma^2 = \Delta e_i)$ and $\widetilde{p_i} \sim N(\mu = p_i, \sigma^2 = \Delta pi.$

For each pair sets $\{\dots, (e_i, \tilde{e}_i), \dots\}$ (Experimental-Experimental), $\{\dots, (p_i, \tilde{p}_i), \dots\}$ (Predicted-Predicted) and $\{, \dots (e_i, \tilde{p}_i), \dots\}$ (Experimental-Predicted) it is possible to determine R^2

(coefficient of determination), *MUE* (mean unsigned error) and *PI* (predictive indices[6]) metrics. Iterating the procedure *n* times yields probability distributions f_{R^2}, f_{MUE} and f_{PI} of these metrics. A confidence interval in the metrics may then be obtained by plotting the resulting probability distributions and choosing lower *a* and upper *b* values of the associated cumulative probability functions *F*. Here probability distributions *f* were computed by re-sampling the sets $n = 10^6$ times, and a 95% confidence interval was chosen setting $a = F^{-1}(0.03)$ and $b = F^{-1}(0.98)$.

Protein-ligand docking protocols

The free energies of binding of the thrombin ligands were also estimated by using the docking software Autodock Vina.[49] This was done to compare the results of the molecular simulations with molecular modelling protocols that do not explicitly consider protein dynamics and hydration. The same protein structure was used for docking calculations. A grid of 20 Å × 20 Å × 20 Å was centered on the thrombin binding site and a docking pose was generated for each ligand. The protein was kept rigid in all docking calculations. In all instances where comparison was possible, it was verified that Vina produced a binding pose very similar to the available experimental data. The Vina predicted binding affinity of the best scoring pose was used for subsequent analyses.

Results and Discussion

Precision and consistency of the computed free energies of binding

Analysis of intermediate results showed that increasing the length of per-λ simulation time from 5 ns to 10 ns did not produce significant changes in the calculated ΔG_{wat}(**A** →**B**) values. Thus statistical uncertainties for these quantities were determined by using block averaging of the data collected with the 10 ns per-λ simulation protocol. To evaluate ΔG_{prot}(**A** →**B**) triplicate runs of 10 ns per-λ each were performed, and the final ΔG_{prot}(**A** →**B**) estimate was taken as the

mean computed free energy change, and uncertainties taken as the standard error of the mean. Average free energy changes ΔG_{wat}(**A**→**B**), ΔG_{prot}(**A**→**B**) and relative binding free energies $\Delta\Delta G_{bind}$ (**A**→**B**) for each perturbation **A**→**B** are reported in Table S1 and S2 and compared with experimental data. With this protocol the maximum uncertainty on the predicted relative binding affinities $\Delta\Delta G_{bind}$ (**A**→**B**) was ca. 0.2-0.4 kcal/mol. The noisier results in the 3 series involved the ligand **3J** and in the 5 series the ligands **5J** and **5I**. The alchemical transformation involving these ligands typically required a high number of atoms to be converted into dummy atoms, e.g. in the transformation **5J** to **5B** 14 atoms were transformed from a fully interacting to a dummy state. Large changes in ligand-excluded volume are associated with larger fluctuations of the free energy gradients.

Consistency and reliability of the computed relative binding free energies was also assessed by considering thermodynamic cycle closures. These are frequently used as quality check for free energy calculations and the discrepancy from zero measures the transformation hysteresis.[50–53] In the selected alchemical transformations for the 3 and 5 series, three and four thermodynamic cycle closures were considered. The analysis showed that the maximum discrepancy for the 3 and 5 series was respectively 0.5±0.2 kcal mol^{-1} (cycle **3B** to **3A** to **3D** in Figure 3A) and -0.9±0.2 kcal mol^{-1} (cycle **5B** to **5A** to **5D** in Figure 3B). The thermodynamic cycles are reported in Figure S1.

Soft-core potentials occasionally introduce serious artefacts

For the perturbation **3K**→**3B** one of the triplicate runs produced a $\Delta G_{prot}(\mathbf{3K} \rightarrow \mathbf{3B})$ value that was significantly different from the other two runs and warranted further investigation. Extending the per-λ simulation time to up to 20 ns (Fig 4A) indicated that the free energy

gradients were all well converged, apart from the gradients at λ = 0.57822 that steadily increases with time. Structural analysis of the computed trajectory at this λ value (Fig 4B) indicates that a pronounced jump in free energy gradients correlates with fusing of a sodium ion with the para carbon atom of the decoupled phenyl ring around t = 2.3 ns. Both particles remain fused for the rest of the simulation. The phenomenon is illustrated with snapshots in Fig 4C. Further analysis indicated that the behavior arose because of a too rapid softening of the Lennard-Jones interactions of the perturbed phenyl ring in **3K.** This enables particles fusion driven by attractive Coulombic interactions between the partial negative charge of the carbon atom and the positive charge of the sodium ion. This effect happened only once due to chance diffusion of the sodium ion near the phenyl ring, and it can be altogether avoided by adjusting soft core parameters to either soften more rapidly with λ the Coulombic interactions, or to soften less rapidly with λ the Lennard-Jones interactions. An alternative could have been to serially turn off partial charges on the perturbed phenyl ring, followed by zeroing of Lennard-Jones parameters. Thus this particular run was discarded and another repeat of 10 ns was performed. No other soft-core related abnormalities were detected in the rest of the dataset.

The free energies of binding are accurate and approach the uncertainties of the experimental data

Table S3 and S4 in the SI reports the relative free energy of binding selecting the ligand **3B** or **5B** as reference ligands respectively. For some ligands the calculation of the relative free energy of binding was calculated by averaging the relative free energy changes obtained along different possible paths in the relative free energy network (Figure 3). For example for the ligand **3E** two

paths were selected: $\mathbf{3B \rightarrow 3D \rightarrow 3E}$ and $\mathbf{3B \rightarrow 3C \rightarrow 3E}$. Statistical errors were propagated accordingly.

Next error analysis was undertaken to assert the reliability and relevance of correlations between measured and computed free energies of binding. The error analysis results are given in Table S5 and Figure 5. The error analysis on the experimental data (Fig 5A) showed that the experimental data for the 5 series is more reliable than the 3 series, which shows a noticeable spread in R^2 and *PI* values. This indicates that uncertainties in the experimental data in the 3 series limit the maximal correlations that may be realistically achieved. The combined data (3 & 5) shows an intermediate behavior between the 3 and 5 Series. The error analysis conducted on the predicted data (Fig 5B) showed that free energies of binding in both series are predicted with similar precision, as evidenced by the high R^2, *PI* and low *MUE* values. The metrics outperform the experimental data and thus uncertainties in the predicted data are unlikely to significantly affect comparison with experimental data. Turning to comparisons between experimental-predicted sets (Fig 5C), it is apparent that both datasets give normally distributed R^2 values with an average value slightly over 0.8 when considered as separate data sets while, when combined, the average R^2 drops below 0.7. The *MUE* values are also normally distributed and the mean *MUE* of the 5 series is lower than for the 3 series. In this case the combined distribution seems to settle between the 3 and 5 series distributions. Interestingly the *PI* distribution is more complex and multimodal in the individual series data sets, but follows approximately a normal distribution for the combined datasets. This indicates that care should be used when assessing uncertainties with this metric since the metric may not be normally or chi-distributed. Figure 6A shows the comparison between experimental and computed data combining the two 3 and 5 series. There is a clear correlation between the data; however the calculated R^2 and *PI* are lower than the same

statistical metrics obtained considering the two data sets separately as shown in Figure 6B and 6C. In Figure 6B the overall agreement is quite satisfactory (R^2 0.81±0.05), and no significant outliers emerge, however, it is apparent that the binding affinity of the larger compounds (e.g. **3J**, **3G**, **3H**, **3K**, **3I**) is systematically overestimated and this is reflected by the slope of the linear regression (1.6 ± 0.3) between experimental and computed free energies of binding. Figure 6C similarly shows that the compounds are well ranked and no outliers emerge. There is also a trend, albeit weaker, for overestimation of the free energies of binding of the larger compounds (**5I**, **5J**, regression slope 1.3±0.2).

Docking energies overall correlate poorly, but show trends within individual series

Docking energies are listed in Table S6. Docking energies between multiple repeats were well reproducible and uncertainties were taken to be ±0.1 kcal mol^{-1}. Comparison of the experimental and docking predicted free energies of binding are shown in Figure 7 along with the determined R^2, *MUE* and *PI* values. Overall the performance on the combined dataset (Figure 7A) is poorer (R^2 0.34±0.07, MUE 0.85±0.04 and *PI* 0.61±0.08) than for the molecular simulation results. Interestingly correlations with experimental data within individual series are noticeably higher (Figure 7B and 7C), although still inferior to the molecular simulation predictions. The main reason for the poor combined scoring is that binding affinities of compounds in the 5 series are systematically underestimated, as evidenced by the low regression slope value of 0.5±0.1.

Free energy calculations capture trends for positive-coupling between X and R^1 substituents.

The computed non-additivity levels for each ligand were next calculated using equation (4). Figure 8 reports the obtained results per ligand calculated for the experimental data, the data

computed with the Sire-OpenMM molecular simulations, and the data computed with the Vina docking calculations. Overall Vina is unable to detect the trend towards positive coupling of the $X = NH_3^+$ and R^1 groups and the results are scattered around the null hypothesis. By contrast the molecular simulations fare much better and statistically significant positive coupling is predicted for 8 of the 9 ligands pairs. The compound pair **3F**/**5F** does not show any statistically significant non-additive effects. It is also apparent that the magnitude of the non-additive effects is systematically weaker than in the experimental data.

The origin of non-additivity of functional group contributions

Why do the molecular simulations predict positive coupling between the $X=NH3^+$ and R^1 groups? Answers were sought via structural analyses of the computed protein-ligand conformational ensembles. Baum et al. previously concluded from B-factor analyses of several thrombin X-ray structures that formation of the charge-assisted hydrogen-bond with Gly216 decreases residual mobility of the R^1 substituents of the 5 series ligands with respect to 3 series ligand. A decrease in residual mobility is not expected to contribute to positive coupling, but could be a consequence of mutually reinforcing interactions between X and R^1 groups. Figure 9 shows that the experimentally observed trend is also observed in the present simulations, but there are two important exceptions. Compound **5A** shows increased flexibility with respect to compound **3A** (positive $<\Delta B_{fac}>$ value). This compound carries the smallest R^1 substituent of the series (R^1=H). Compound **5F** does not show any change in residual mobility with respect to compound **3F**. Interestingly **5F** is the only compound with no significant computed non-additive level (Figure 8). Figure 9 also shows that reduction in ligand mobility correlates with decreases

in average distances between the nitrogen atom of the $X=NH_3^+$ group, and the backbone carbonyl of Gly216. Compounds **5A** and **5F** are again the two exceptions.

Different hypotheses about the origin of non-additivity have been previously put forward. One can speculate that the extra-hydrogen bond with the backbone carbonyl of Gly216 holds the R^1 substituents in conformations that pack better with the S3 pocket side-chains.[54] Alternatively the larger R^1 substituents push the ligand slightly away from the S3 pocket, causing a shortening of the hydrogen-bond distances between the extra amino group of the 5 series ligands and the backbone carbonyl of Gly216, and one of the ligands amide carbonyl and Gly216 backbone NH.[17]

Clarification was sought by extracting the contribution of protein-ligand interaction energies to the non-additivity levels from the computed trajectories (equation 6). The thrombin binding site was partitioned in four groups that contact different sections of the ligands (Figure 10A). Figure 10B shows that there is no strong correlation between changes in ligand interaction energies with S3/S4 pocket residues and computed non-additivity levels. **5K** is the only ligand that shows weak but statistically significant enhanced lipophilic contacts. Figure 10B also shows that all 5 series ligands apart from **5A**, **5F** and **5J** exhibit enhanced interactions with Gly216. This was expected for **5A** and **5F** in light of the changes in average hydrogen-bond distances shown in Figure 9. For **5J** favorable contributions of protein-ligand interaction energies to the computed non-additivity level cannot be detected owing to high statistical errors. This was already reflected by the relatively high uncertainties in the computed non-additivity level for this compound (Figure 8). Remarkably, Figure 10B also shows that ligand **5F** is the only compound that exhibits enhanced interactions with S2 pocket residues His57, Tyr60A and Trp60D. These interactions occur via packing of the ligand pyrrolidine ring. All other ligands show a trend for

weakened interactions with this group of protein residues, but the trend is significant only for **5E**, **5H** and **5I**. Lastly Figure 10B also shows that contacts with residues in the S1 pocket also contribute favorably to the binding of **5H**, **5I** but disfavor **5F**. That is, the trend is the reverse of what was observed for interactions with S2 pocket residues.

The picture that emerges is that anchoring of the 5 series ligands via formation of the hydrogen bond between the X group and Gly216 does not strengthen lipophilic contact between R^1 groups and the S3/S4 pocket residues. However, an increase in the size of the R^1 substituent does generally contribute to an energetically favorable shortening of hydrogen-bonding distances between the X group and the backbone carbonyl of Gly216. Nevertheless, it is misleading to ascribe all non-additivity to this single interaction since weakened interactions are also observed with S2 pocket residues, and enhanced interactions are observed with S1 pocket residues. For compound **5F**, the opposite happens, that is weakened interactions with S1, S3/S4 pockets residues, and enhanced interactions with S2 pocket residues. This is achieved by small readjustments of **5F** and binding site residues, which is sufficient to abrogate positive coupling between X =NH_3^+ and R^1=i-Butyl groups.

For **5A** the above explanation is unsatisfactory since this compound lacks any substituent interacting with the S3 pocket, does not exhibit statistically significant changes in protein-ligand interaction energies (Figure 10B), and yet shows significant non-additivity in both experiments and simulations (Figure 8). The data in Figure 9 suggests that enhanced residual mobility of **5A** when bound to thrombin may contribute to non-additivity. More quantitative statements about entropic contributions to non-additivity for **5A** were not pursued owing to the difficulties encountered in accurately evaluating entropic changes and partitioning entropy into meaningful components.[18] The analysis presented in Figure 10 could have been extended to other

components of the potential energy function, for instance changes in hydration enthalpies have been proposed to contribute to non-additivity in some protein-ligand complexes,[55] and these could be computed using molecular simulation trajectory analyses methodologies.[56–58] However, exploratory calculations suggest that converging hydration energy components to a level of precision sufficient to enable meaningful comparison between ligands will require significantly longer simulations than what was achieved here. Nevertheless it is now clear from the data depicted in Figure 10B that a profound understanding of non-additive effects in this system requires consideration of interactions occurring across the entire thrombin binding site.

Conclusion

In summary the alchemical free energy calculation protocols used here could accurately predict the relative free energies of binding of two congeneric series of thrombin ligands. Evidence for statistically significant non-additivity of interaction energies between hydrophobic groups R^1 and the polar group $X=NH_3^+$ was obtained, which required aggregated molecular dynamics sampling time of ca. 10 μs. The experimentally observed trend for positive coupling between these two groups was reproduced for all compounds but **5F**, albeit with a weaker magnitude. Further work is needed to establish why **5F** is an outlier. This could be pursued for instance by evaluating whether the results for **5F** are an artefact of the particular ligand force-field use here.

Although docking protocols were also successful at correlating docking energies with binding affinities within individual ligand series, a much poorer correlation was obtained when ranking both series together. This is because the docking protocol was unable to capture any coupling between interactions of the R^1 and X groups. This suggests that non-additive effects in this

dataset originate from changes in solvation and/or dynamics, which are accounted for explicitly in the molecular simulation, but largely lacking in docking calculations.

Analysis of the computed molecular simulation trajectories suggests that changes in contacts between R^1 groups and S3/S4 pocket residues is not a major contribution to the observed positive coupling. However, shortening of hydrogen-bonding distances between the ligands and Gly216 is an important source of non-additivity, as was suggested earlier.[17] Yet changes in this interaction do not explain all observed non-additivity levels, and other changes in interactions with protein residues occur across the entire binding site. Therefore coupling between the $X=NH_3^+$ and R^1 substituents is the result of an energetic balance between numerous interactions, whose final outcome depends non-trivially on the precise chemical nature of R^1.

Should we be surprised? While we ultimately hope for models that produce simple actionable ligand design guidelines, Nature is not concerned with the partitioning of protein-ligand interactions into components that satisfy human intuition. While there may not be a simple explanation to the origin of non-additivity of protein-ligand interactions for the compounds investigated here, it is gratifying that modern potential energy functions and molecular simulation protocols have achieved the precision and accuracy needed to capture this phenomenon.

ASSOCIATED CONTENT

Supporting Information.

Table of estimated ΔG_{wat}(**A→B**), ΔG_{prot}(**A→B**) and $\Delta\Delta G_{bind}$ (**A→B**) values for each perturbation. Cycle closure errors. Table of estimated free energies of binding relative to reference compounds **3B** and **5B**. Table of uncertainties for the dataset correlation metrics. Table

of docking predicted binding free energies. Instructions to download all input files for the Sire-OpenMM calculations.

AUTHOR INFORMATION

Corresponding Author

* E-mail: mail@julienmichel.net Phone number: +44 (0)131 650 4797

Current Address:

‖ Gaetano Calabrò - Pharmaceutical Department, University of California Irvine, Irvine, CA 92697 USA.

‖ Francis Powlesland - IBM 76/78 Upper Ground, London, SE1 9PZ, UK

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

J.M. is supported by a Royal Society University Research Fellowship. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement No. 336289. This research was supported by a NVIDIA academic partnership. Parts of the simulations were performed using the Supercomputing Service Blue Crystal at the University of Bristol (https://www.acrc.bris.ac.uk). This research was supported by CCP-BioSim via travel grants to G. C. CCP-BioSim is funded by the EPSRC (EP/J010588/1). C. J. W. acknowledges support of

BrisSynBio, a BBSRC/EPSRC Synthetic Biology Research Centre (BB/L01386X/1) and the BBSRC (BB/K016601/1).

References

(1) Heald, R.; Bowman, K. K.; Bryan, M. C.; Burdick, D.; Chan, B.; Chan, E.; Chen, Y.; Clausen, S.; Dominguez-Fernandez, B.; Eigenbrot, C.; et al. Noncovalent Mutant Selective Epidermal Growth Factor Receptor Inhibitors: A Lead Optimization Case Study. *J. Med. Chem.* **2015**, *58* (22), 8877–8895.

(2) Smith, A. J. T.; Zhang, X.; Leach, A. G.; Houk, K. N. Beyond Picomolar Affinities: Quantitative Aspects of Noncovalent and Covalent Binding of Drugs to Proteins. *J. Med. Chem.* **2009**, *52* (2), 225–233.

(3) Whitesides, G. M.; Krishnamurthy, V. M. Designing Ligands to Bind Proteins. *Q. Rev. Biophys.* **2005**, *38* (4), 385–395.

(4) Bissantz, C.; Kuhn, B.; Stahl, M. A Medicinal Chemist's Guide to Molecular Interactions. *J. Med. Chem.* **2010**, *53* (14), 5061–5084.

(5) Freire, E. Do Enthalpy and Entropy Distinguish First in Class from Best in Class? *Drug Discov. Today* **2008**, *13* (19-20), 869–874.

(6) Ruben, A. J.; Kiso, Y.; Freire, E. Overcoming Roadblocks in Lead Optimization: A Thermodynamic Perspective. *Chem. Biol. Drug Des.* **2006**, *67* (1), 2–4.

(7) Chaires, J. B. Calorimetry and Thermodynamics in Drug Design. *Annu. Rev. Biophys.* **2008**, *37* (1), 135–151.

(8) Bravi, G.; Gancia, E.; Green, D. V. S.; Hann, M. M. Modelling Structure-Activity Relationships. In *Virtual Screening for Bioactive Molecules*; Böhm, H.-J., Schneider, G., Eds.; Wiley-VCH Verlag GmbH, 2000; pp 81–116.

(9) Agrafiotis, D. K.; Bandyopadhyay, D.; Wegner, J. K.; van Vlijmen, H. Recent Advances in Chemoinformatics. *J. Chem. Inf. Model.* **2007**, *47* (4), 1279–1293.

(10) Kenny, P. W.; Sadowski, J. Structure Modification in Chemical Databases. In *Chemoinformatics in Drug Discovery*; Oprea, T. I., Ed.; Wiley-VCH Verlag GmbH & Co. KGaA, 2005; pp 271–285.

(11) Free, S. M.; Wilson, J. W. A Mathematical Contribution to Structure-Activity Studies. *J. Med. Chem.* **1964**, *7* (4), 395–399.

(12) Hansch, C. Quantitative Approach to Biochemical Structure-Activity Relationships. *Acc. Chem. Res.* **1969**, *2* (8), 232–239.

(13) Mark, A. Decomposition of the Free Energy of a System in Terms of Specific Interactions Implications for Theoretical and Experimental Studies. *J. Mol. Biol.* **1994**, *240* (2), 167–176.

(14) Dill, K. A. Additivity Principles in Biochemistry. *J. Biol. Chem.* **1997**, *272* (2), 701–704.

(15) van Gunstener, W. F.; Beutler, T. C.; Fraternali, F.; King, P. M.; Mark, A. E.; Smith, P. E. *Computation of Free Energy in Practice: Choice of Approximations and Accuracy Limiting Factors*; Computer Simulation of Biomolecular Systems, theoretical and experimental applications.; Escom Science, 1993; Vol. 2.

(16) Patel, Y.; Gillet, V. J.; Howe, T.; Pastor, J.; Oyarzabal, J.; Willett, P. Assessment of Additive/Nonadditive Effects in Structure−Activity Relationships: Implications for Iterative Drug Design. *J. Med. Chem.* **2008**, *51* (23), 7552–7562.
(17) Baum, B.; Muley, L.; Smolinski, M.; Heine, A.; Hangauer, D.; Klebe, G. Non-Additivity of Functional Group Contributions in Protein–Ligand Binding: A Comprehensive Study by Crystallography and Isothermal Titration Calorimetry. *J. Mol. Biol.* **2010**, *397* (4), 1042–1054.
(18) Michel, J. Current and Emerging Opportunities for Molecular Simulations in Structure-Based Drug Design. *Phys. Chem. Chem. Phys.* **2014**, *16* (10), 4465–4477.
(19) Woods, C. J.; Malaisree, M.; Michel, J.; Long, B.; McIntosh-Smith, S.; Mulholland, A. J. Rapid Decomposition and Visualisation of Protein–ligand Binding Free Energies by Residue and by Water. *Faraday Discuss.* **2014**, *169* (0), 477–499.
(20) *Schrödinger Release 2011-1: Maestro, Version 9.2, Schrödinger, LLC, New York, NY, 2011*.
(21) Ko, J.; Dongseon Lee, H. P.; Evangelos Coutsias, J. L.; Seok, C. The FALC-Loop Web Server for Protein Loop Modeling. *Nucleic Acids Res.* **2011**, *39* (2), 210–214.
(22) Loeffler, H. H.; Michel, J.; Woods, C. FESetup: Automating Setup for Alchemical Free Energy Simulations. *J. Chem. Inf. Model.* **2015**, *55* (12), 2485–2490.
(23) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2004**, *25* (9), 1157–1174.
(24) Jakalian, A.; Bush, B. L.; Jack, D. B.; Bayly, C. I. Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: I. Method. *J. Comput. Chem.* **2000**, *21* (2), 132–146.
(25) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comput. Chem.* **2005**, *26* (1), 114–114.
(26) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79* (2), 926–935.
(27) Aaqvist, J. Ion-Water Interaction Potentials Derived from Free Energy Perturbation Simulations. *J. Phys. Chem.* **1990**, *94* (21), 8021–8024.
(28) Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters. *Proteins* **2006**, *65* (3), 712–725.
(29) Michel, J.; Essex, J. Prediction of Protein-Ligand Binding Affinity by Free Energy Simulations: Assumptions, Pitfalls and Expectations. *J. Comput. Aided Mol. Des.* **2010**, *24* (8), 639–658.
(30) Michel, J.; Foloppe, N.; Essex, J. W. Rigorous Free Energy Calculations in Structure-Based Drug Design. *Mol. Inform.* **2010**, *29* (8-9), 570–578.
(31) Woods, C. J.; Calabro, G.; Michel, J. *Sire Molecular Simulation Framework, Revision 2270, Http://siremol.org/ (Accessed May 19, 2016)*.
(32) Eastman, P.; Pande, V. S. OpenMM: A Hardware-Independent Framework for Molecular Simulations. *Comput. Sci. Eng.* **2010**, *12* (4), 34–39.
(33) Rovigatti, L.; Šulc, P.; Reguly, I. Z.; Romano, F. A Comparison between Parallelization Approaches in Molecular Dynamics Simulations on GPUs. *J. Comput. Chem.* **2015**, *36* (1), 1–8.

(34) Davis, J. E.; Ozsoy, A.; Patel, S.; Taufer, M. Towards Large-Scale Molecular Dynamics Simulations on Graphics Processors. In *Bioinformatics and Computational Biology*; Rajasekaran, S., Ed.; Lecture Notes in Computer Science; Springer Berlin Heidelberg, 2009; pp 176–186.
(35) Xu, D.; Williamson, M. J.; Walker, R. C. Advancements in Molecular Dynamics Simulations of Biomolecules on Graphical Processing Units. In *Annual Reports in Computational Chemistry*; Wheeler, R. A., Ed.; Elsevier, 2010; Vol. 6, pp 2–19.
(36) Shyu, C.; Ytreberg, M. Reducing the Bias and Uncertainty of Free Energy Estimates by Using Regression to Fit Thermodynamic Integration Data. *J. Comput. Chem.* **2009**, *30* (14), 2297–2304.
(37) Michel, J.; Verdonk, M.; Essex, J. Protein-Ligand Complexes: Computation of the Relative Free Energy of Different Scaffolds and Binding Modes. *J Chem Theory Comput* **2007**, *3*, 1645–1655.
(38) Kirkwood, J. Statistical Mechanics of Fluid Mixtures. *J. Chem. Phys.* **1935**, *3* (5), 300–313.
(39) Kirkwood, J. G. *Theory of Liquids*; Gordon and Breach, 1968.
(40) Mezei, M. The Finite Difference Thermodynamic Integration, Tested on Calculating the Hydration Free Energy Difference between Acetone and Dimethylamine in Water. *J. Chem. Phys.* **1987**, *86* (12), 7084–7088.
(41) Zwanzig, R. High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases. *J. Chem. Phys.* **1954**, *22* (8), 1420–1426.
(42) Chow, K.-H.; Ferguson, D. M. Isothermal-Isobaric Molecular Dynamics Simulations with Monte Carlo Volume Sampling. *Comput. Phys. Commun.* **1995**, *91* (1–3), 283–289.
(43) Åqvist, J.; Wennerström, P.; Nervall, M.; Bjelic, S.; Brandsdal, B. O. Molecular Dynamics Simulations of Water and Biomolecules with a Monte Carlo Constant Pressure Algorithm. *Chem. Phys. Lett.* **2004**, *384* (4–6), 288–294.
(44) Andersen, H. C. Molecular Dynamics Simulations at Constant Pressure And/or Temperature. *J. Chem. Phys.* **1980**, *72* (4), 2384–2393.
(45) Barker, J. A.; Watts, R. O. Monte Carlo Studies of the Dielectric Properties of Water-like Models. *Mol. Phys.* **1973**, *26* (3), 789–792.
(46) Fernández, D. P.; Mulev, Y.; Goodwin, A. R. H.; Sengers, J. M. H. L. A Database for the Static Dielectric Constant of Water and Steam. *J. Phys. Chem. Ref. Data* **1995**, *24* (1), 33–70.
(47) Brown, S. P.; Muchmore, S. W.; Hajduk, P. J. Healthy Skepticism: Assessing Realistic Model Performance. *Drug Discov. Today* **2009**, *14* (7-8), 420–427.
(48) Luccarelli, J.; Michel, J.; Tirado-Rives, J.; Jorgensen, W. L. Effects of Water Placement on Predictions of Binding Affinities for p38α MAP Kinase Inhibitors. *J. Chem. Theory Comput.* **2010**, *6* (12), 3850–3856.
(49) Trott, O.; Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2010**, *31* (2), 455–461.
(50) Chipot, C. *Free Energy Calculations: Theory and Applications in Chemistry and Biology (Springer Series in Chemical Physics)*; Springer, 2007.
(51) Michel, J.; Tirado-Rives, J.; Jorgensen, W. L. Energetics of Displacing Water Molecules from Protein Binding Sites: Consequences for Ligand Optimization. *J. Am. Chem. Soc.* **2009**, *131* (42), 15403–15411.

(52) Michel, J.; Verdonk, M. L.; Essex, J. W. Protein-Ligand Binding Affinity Predictions by Implicit Solvent Simulations: A Tool for Lead Optimization? *J. Med. Chem.* **2006**, *49* (25), 7427–7439.
(53) Mishra, S. K.; Calabró, G.; Loeffler, H. H.; Michel, J.; Koča, J. Evaluation of Selected Classical Force Fields for Alchemical Binding Free Energy Calculations of Protein-Carbohydrate Complexes. *J. Chem. Theory Comput.* **2015**, *11* (7), 3333–3345.
(54) Muley, L.; Baum, B.; Smolinski, M.; Freindorf, M.; Heine, A.; Klebe, G.; Hangauer, D. G. Enhancement of Hydrophobic Interactions and Hydrogen Bond Strength by Cooperativity: Synthesis, Modeling, and Molecular Dynamics Simulations of a Congeneric Series of Thrombin Inhibitors. *J. Med. Chem.* **2010**, *53* (5), 2126–2135.
(55) Biela, A.; Betz, M.; Heine, A.; Klebe, G. Water Makes the Difference: Rearrangement of Water Solvation Layer Triggers Non-Additivity of Functional Group Contributions in Protein–Ligand Binding. *ChemMedChem* **2012**, *7* (8), 1423–1434.
(56) Gerogiokas, G.; Calabro, G.; Henchman, R. H.; Southey, M. W. Y.; Law, R. J.; Michel, J. Prediction of Small Molecule Hydration Thermodynamics with Grid Cell Theory. *J. Chem. Theory Comput.* **2014**, *10* (1), 35–48.
(57) Michel, J.; Henchman, R. H.; Gerogiokas, G.; Southey, M. W. Y.; Mazanetz, M. P.; Law, R. J. Evaluation of Host–Guest Binding Thermodynamics of Model Cavities with Grid Cell Theory. *J. Chem. Theory Comput.* **2014**, *10* (9), 4055–4068.
(58) Gerogiokas, G.; Southey, M. W. Y.; Mazanetz, M. P.; Hefeitz, A.; Bodkin, M.; Law, R. J.; Michel, J. Evaluation of Water Displacement Energetics in Protein Binding Sites with Grid Cell Theory. *Phys. Chem. Chem. Phys.* **2015**, *17* (13), 8416–8426.

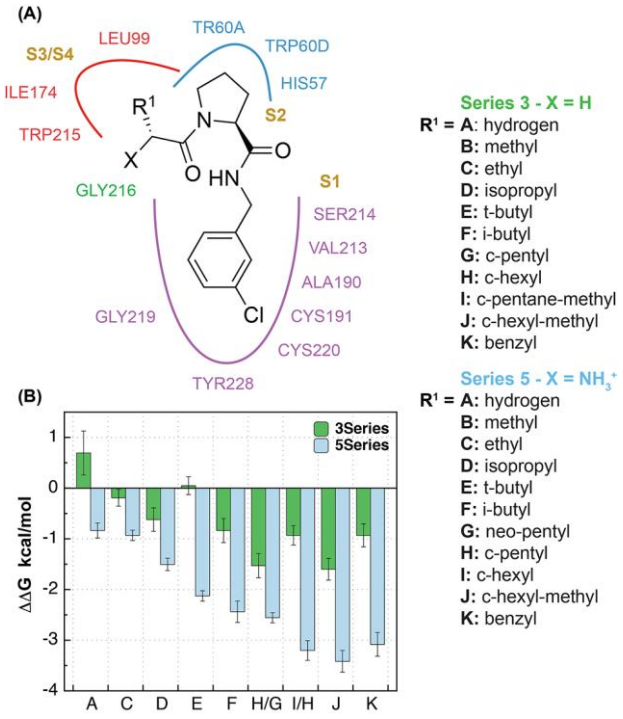


Figure 1. A) Depiction of the thrombin-ligand complexes considered in this study. The main protein residues involved in interactions with the ligands are shown. Figure adapted from Baum et al.[17] (B) Histogram of the experimental free energy of binding relative to compound **B** for series 3 and series 5 compounds with equivalent R^1 groups. It is apparent that the relative free energy of binding of series 5 ligands is more negative than for series 3 ligands. The ligand names are those used by Baum et al. [17] paper. Ligand **G** in the 3 series maps to ligand **H** in the 5 series, while ligand **H** in the 3 series maps to ligand **I** in the 5 series.

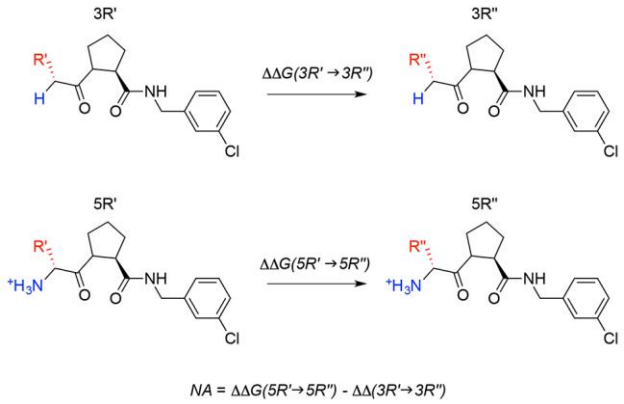


Figure 2. The ligand 5R'' presents both the modifications of the ligands 5R' and 3R''. Therefore, the non-additivity *NA* can be evaluated calculating the difference between the relative free energy of binding $\Delta\Delta G(5R'\rightarrow 5R'')$ and $\Delta\Delta G(3R'\rightarrow 3R'')$.

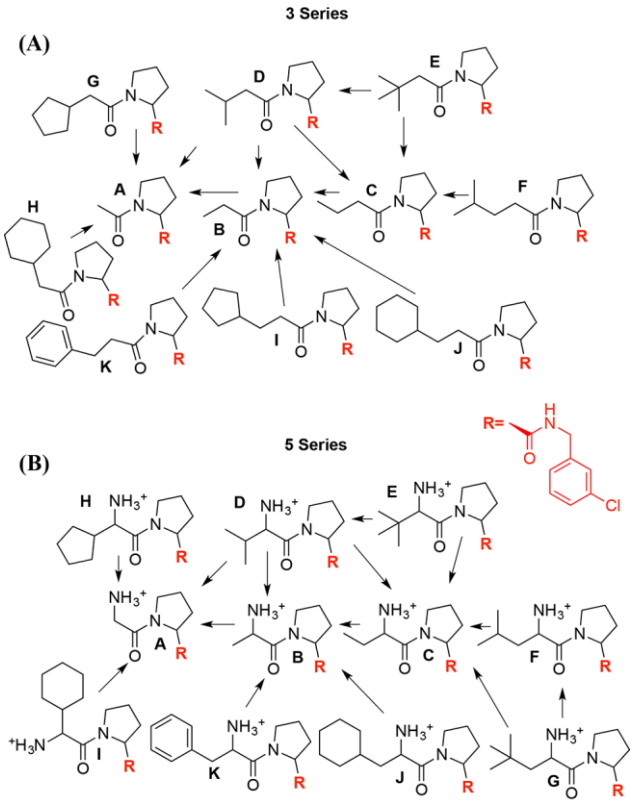


Figure 3. The network of relative free energy calculations map used to determine the free energy of binding of A) series 3 and B) series 5 ligands with respect to reference compounds **3B** and **5B**.

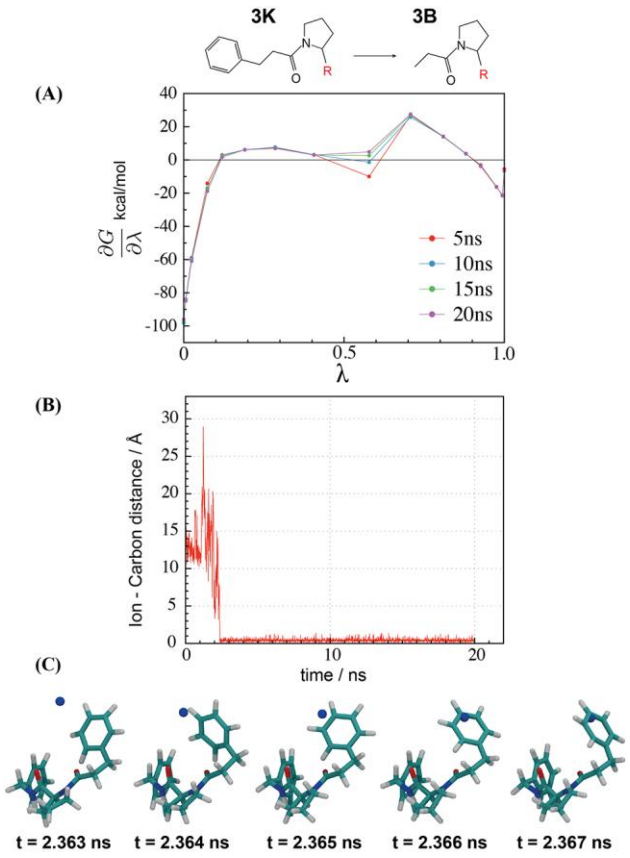


Figure 4. The perturbation **3K→3B** presented unusual convergence behavior in one of the triplicate runs of the protein-ligand complex. A) Extending the per-λ duration of the simulations from 5 ns to up to 20 ns demonstrated high convergence of free energy gradients apart from λ = 0.57822 where the gradients steadily increase with time. B) A plot of the distance between one sodium Na+ and the para carbon atom of the decoupled phenyl ring indicates that the ion overlaps with the carbon atom after ca. 2.3 ns, and this correlates with a jump in the free energy gradient. C) Snapshots sampled from the simulation around 2.3 ns.

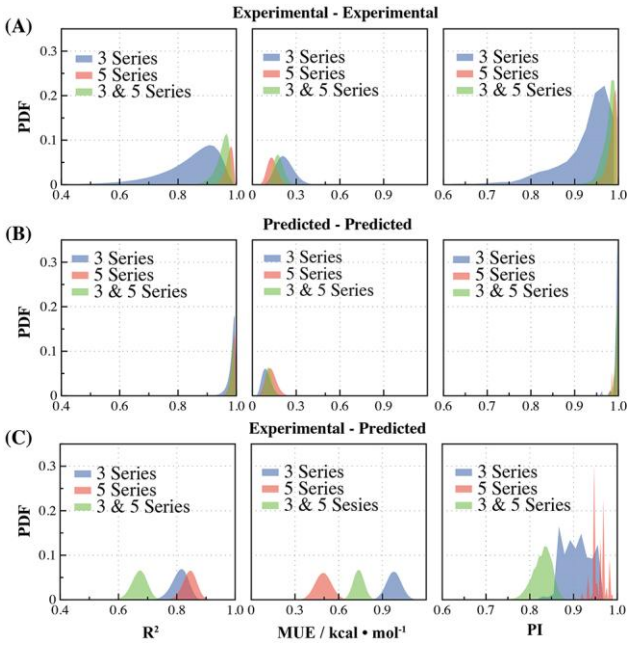


Figure 5. The computed Probability Distribution Functions (PDF) for R^2, MUE and PI metrics of the 3, 5 Series and combining them (3 & 5 Series). A) Experimental-Experimental pairing. B) Predicted-Predicted pairing. C) Experimental-Predicted pairing.

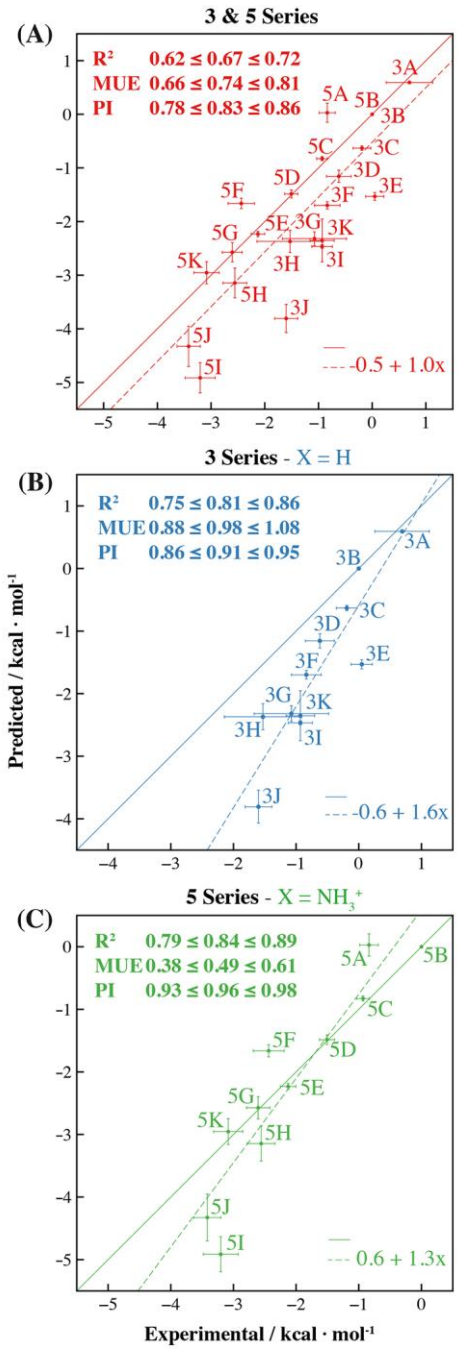


Figure 6. Comparison of measured and molecular simulations predicted relative free energies of binding. A) 3 & 5 Series combined. B) 3 Series and C) 5 Series. The free energies of binding are relative to ligand 3B or 5B. R^2, MUE and PI metrics and uncertainty intervals are also plotted, along with a linear regression slope and intercept (dashed-line). The solid line has slope 1 and intercept 0.

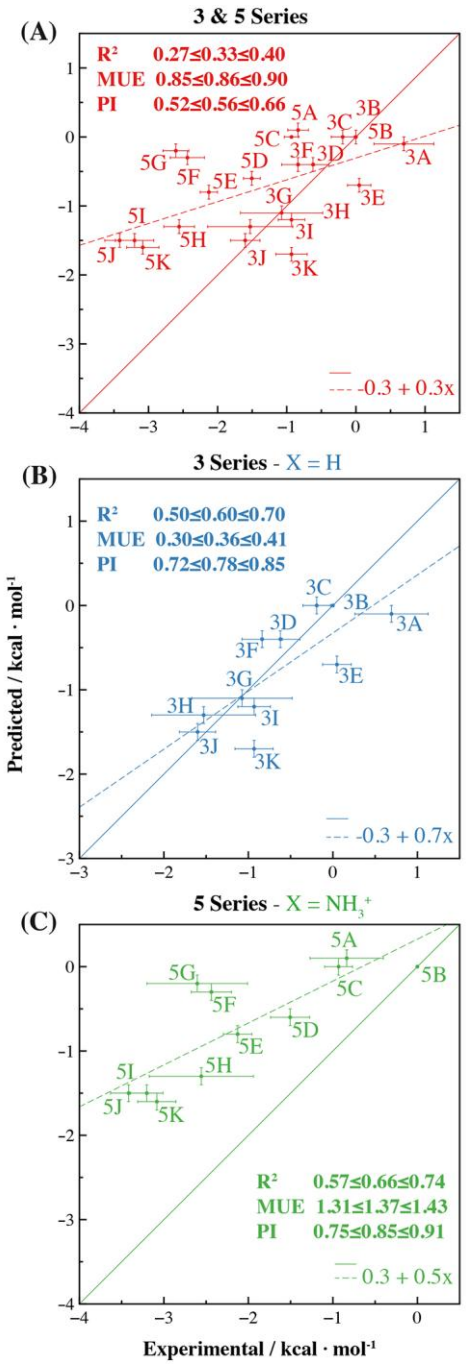


Figure 7. Comparison of measured and docking predicted relative free energies of binding. A) 3 & 5 Series combined. B) 3 Series and C) 5 Series. The free energies of binding are relative to ligand 3B or 5B

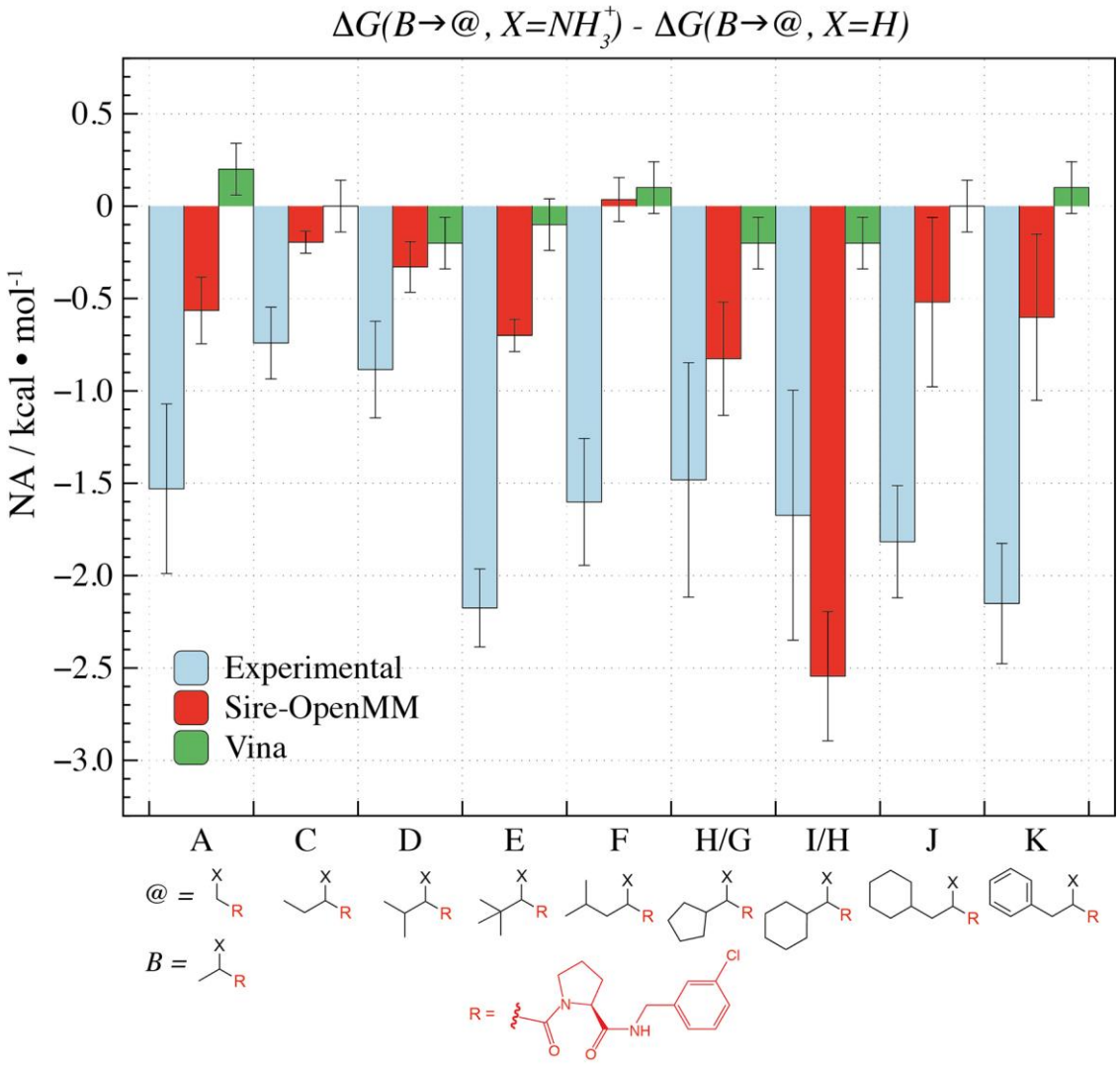


Figure 8. Comparison of non-additivity levels of the relative free energies of binding for experimental data (blue), Sire-OpenMM predicted free energies of binding (red), and Vina predicted free energies of binding (green). Note that ligands **G** and **H** in 3 series map respectively to ligands **H** and **I** of series 5.

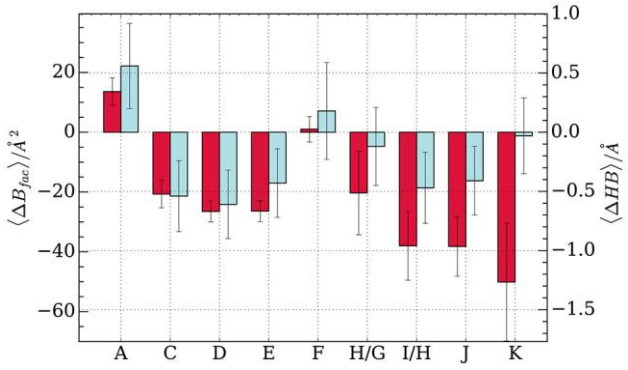


Figure 9. Changes in average B-factors of R^1 atoms of 5 series ligand with respect to 3 series ligands (red), and changes in average hydrogen-bond distance between X=NH_3+ and the backbone carbonyl of Gly216 of 5 series ligands with respect to ligand **5B** (blue). Error bars denote one standard error of the mean.

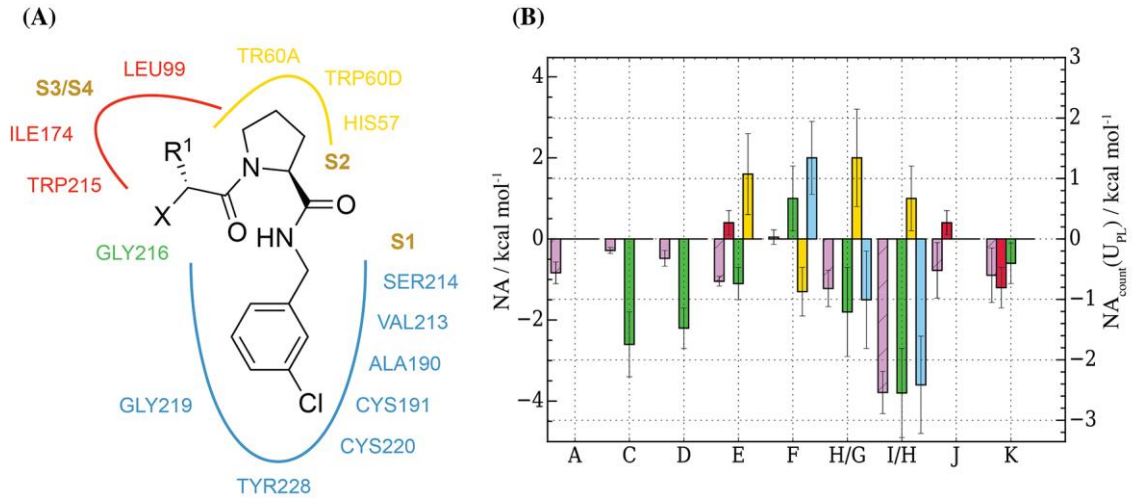
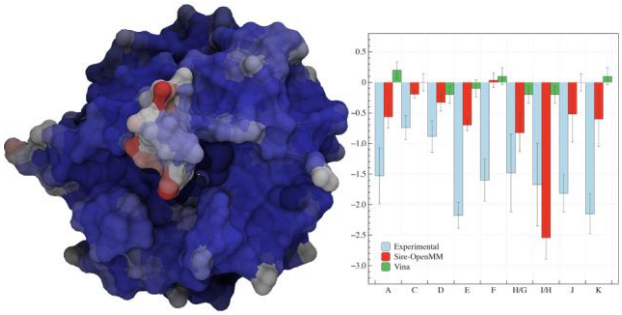


Figure 10. Contribution of protein-ligand interaction energy components to non-additivity levels. A) The thrombin binding site was decomposed in four regions. B) The computed non additivity levels are represented in pink. Other histograms indicate the contributions of protein-ligand interactions energies to non-additivity levels from the S3/S4 pocket group (red), from Gly216 (green), from the S2 pocket group (yellow), S1 pocket group (light blue). Error bars denote one standard error of the mean. For clarity only contributions that deviate from 0 kcal mol^{-1} by at least one standard error are shown.


0.5
0
−0.5
−1.0
−1.5
−2.0
−2.5
−3.0
Experimental
Sire-OpenMM
Vina
A
C
D
E
F
H/G
I/H
J
K

TOC GRAPHIC